Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 9 (2012) 764 – 773

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2012

Knowledge-Based Response Correction and Adaptive Design Specifications for Microwave Design Optimization

Slawomir Koziel[a*], Stanislav Ogurtsov[a], and Leifur Leifsson[a]

[a]*Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland*

Abstract

Simulation-based optimization has become an important design tool in microwave engineering. Yet, employing electromagnetic (EM) solvers in the design process is a challenging task, primarily due to a high-computational cost of an accurate EM simulation. This paper is focused on efficient EM-driven design optimization techniques that utilize physically-based low-fidelity models, normally based on coarse-discretization EM simulations. The presented methods attempt to exploit as much of the knowledge about the system or device of interest embedded in the low-fidelity model as possible, so as to reduce the computational cost of the design process. Unlike many other surrogate-based approaches, the techniques discussed here are non-parametric ones, i.e., they are not based on analytical formulas. The paper presents several specific methods, including those based on correcting the low-fidelity model response (adaptive response correction and shape-preserving response prediction), as well as on suitable modification of the design specifications. Formulations, application examples and the discussion of advantages and disadvantages of these techniques are also included.

Keywords: Simulation-driven design, microwave engineering, electromagnetic simulation, antenna design, surrogate models.

1. Introduction

Electromagnetic (EM) simulation has become an inherent part of contemporary microwave design process. It is used for verification purposes, but also, more and more often, to adjust the designable parameters of the structure under consideration so that given performance specifications are met. The EM-simulation-driven design is actually a must for the growing number of devices, where systematic design procedures or design-ready theoretical models are not sufficiently accurate, including substrate-integrated circuits [1], ultrawideband antennas [2] or dielectric resonator antennas [3]. EM-based design optimization is a challenging task though. Probably the biggest obstacle is the high computational cost of accurate full-wave simulation, which makes the use of conventional optimization algorithms prohibitive as these techniques require a large numbers of EM simulations. Another issue is the numerical noise that is inherent to EM simulation, particularly if the solver uses adaptive meshing. These problems can be alleviated to some extent by using adjoint sensitivities [4] that have recently become available in some major commercial EM solvers [5], [6]. In general, robust automated and computationally efficient EM-based optimization is still an open problem. In practice, the most common simulation-driven design approach is to perform parameter sweeps guided by engineering experience. This is, however, tedious and does not guarantee optimal results.

* Corresponding author. Tel.: +354-599-6376 ; fax: +354-599-6489; *E-mail address*: koziel@ru.is

doi:10.1016/j.procs.2012.04.082

A number of techniques for modeling and simulation-driven design of microwave structures have emerged over the recent years including the methods that exploit artificial neural networks [7], [8], fuzzy systems [9], or kriging [10], as well as surrogate-based techniques such as space mapping (SM) [11]-[16], simulation-based tuning [17], [18], and combination of both [19], [20]. The last three approaches offer computationally efficient design optimization where, under certain circumstances, a satisfactory design can be obtained after a few high-fidelity (or fine) EM simulations of the structure of interest [11]. SM and tuning SM (TSM) [19], exploit physically-based low-fidelity (or coarse) model that is a basis to build a surrogate model – computationally cheap and yet reasonably accurate representation of the fine model. The surrogate is iteratively updated and re-optimized instead of optimizing the fine model directly [11]. SM and TSM are both potentially very efficient, but these methods may not be straightforward to implement and automate [12], which, so far, is a limitation for their widespread use by the researchers and designers.

Probably the simplest way of exploiting physically-based coarse model to perform efficient design optimization is through correction of the coarse model response in order to obtain zero- and (in some cases) first-order consistency [21] between the surrogate and the fine model. The corrected coarse model becomes a reliable prediction tool that can be used to find approximate optimum of the fine model. Several methods exploiting this approach have been proposed in microwave engineering recently including manifold mapping (MM) [22] and multi-point response correction [23]. Formally, output SM [11] and its variations [23] also belong to this group, as well as the neuro-space mapping techniques [8]. All of these techniques can be categorized as parametric response correction methods, because the adjustment of the low-fidelity model is realized by means of explicit formulas, parameters of which can be calculated using available low- and high-fidelity model data.

This paper is focused on non-parametric approaches that allow us to fully exploit the knowledge about the structure of interest embedded in the low-fidelity model. The techniques such as adaptive response correction (ARC) [24] and shape-preserving response prediction (SPRP) [25] do not use explicit parameters. They utilize relationships between the low- and high-fidelity model responses as functions of the design variables, not just their value at specific designs, which normally results in much better generalization capability when compared to parametric approaches. The adaptively adjusted design specifications (AADS) technique [26] is based on a similar idea. In AADS, however, the relationship between the low- and high-fidelity models is reflected in proper modification of the design specifications without directly adjusting the low-fidelity model itself, which makes this method extremely simple to implement and yet robust.

Here, we recall the response correction concept as well as the formulation of the optimization algorithm exploiting a corrected low-fidelity model. We focus on the non-parametric correction techniques, however, for the sake of the paper being self-contained, parametric approaches are recalled as well. Additionally, we discuss the adaptively adjusted design specification method. All the techniques presented here are illustrated using representative microwave design examples.

2. Design Optimization Using Response-Corrected Physically-Based Coarse Models

The design optimization task can be formulated as a following minimization problem

$$\boldsymbol{x}_f^* \in \arg\min_{\boldsymbol{x}} U\left(\boldsymbol{R}_f(\boldsymbol{x})\right) \tag{1}$$

where $\boldsymbol{R}_f \in R^m$ denotes the response vector of a fine model of the device of interest evaluated through high-fidelity EM simulation; $\boldsymbol{x}$ is a vector of design parameters. The response $\boldsymbol{R}_f(\boldsymbol{x})$ might be, e.g., the modulus of the reflection coefficient $|S_{11}|$ evaluated at m different frequencies. U is a given scalar merit function, e.g., a norm, or a minimax function with upper and lower specifications; $\boldsymbol{x}_f^*$ is the optimal design to be determined. It is assumed that $\boldsymbol{R}_f$ is evaluated through computationally expensive EM simulation so that solving (1) directly may be impractical.

Let $\boldsymbol{R}_c(\boldsymbol{x}) \in R^m$ be the response vector of the coarse model of the structure of interest, i.e., computationally cheap but less accurate representation of the fine model. It is also assumed that the coarse model is physically based so that it "embeds" some knowledge about the fine model $\boldsymbol{R}_f$. Typical examples include equivalent circuits or coarse-discretization EM models. The coarse model can be used to construct a surrogate $\boldsymbol{R}_s$ of the fine model so that an approximate solution to (1) can be obtained by optimizing $\boldsymbol{R}_s$ instead of $\boldsymbol{R}_f$. In particular, the surrogate can be obtained as follows:

$$\boldsymbol{R}_s(\boldsymbol{x}) = \boldsymbol{C}(\boldsymbol{R}_c(\boldsymbol{x})) \tag{2}$$

where $\boldsymbol{C} : R^m \to R^m$ is a response correction function.

Typically, the optimization process involving surrogates is performed in an iterative manner so that the sequence $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$, ..., of (hopefully) better and better approximations to $\boldsymbol{x}_f^*$ are obtained as

$$\boldsymbol{x}^{(i+1)} = \arg\min_{\boldsymbol{x}} U\left(\boldsymbol{R}_s^{(i)}(\boldsymbol{x})\right) \tag{3}$$

with $\boldsymbol{R}_s^{(i)}$ being the surrogate model at iteration i. Here, $\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{C}^{(i)}(\boldsymbol{R}_c(\boldsymbol{x}))$, where $\boldsymbol{C}^{(i)}$ is the correction function at iteration i. For surrogates constructed using response correction, we typically request that at least zero-order consistency between the surrogate and the fine model is satisfied, i.e., $\boldsymbol{R}_s^{(i)}(\boldsymbol{x}^{(i)}) = \boldsymbol{R}_f(\boldsymbol{x}^{(i)})$. It can be shown [27] that satisfaction of first-order consistency, i.e., $\boldsymbol{J}[\boldsymbol{R}_s^{(i)}(\boldsymbol{x}^{(i)})] = \boldsymbol{J}[\boldsymbol{R}_f(\boldsymbol{x}^{(i)})]$ (here, $\boldsymbol{J}[\cdot]$ denotes the Jacobian of the respective model), guarantees convergence of $\{\boldsymbol{x}^{(i)}\}$ to a local optimum of $\boldsymbol{R}_f$ assuming that (3) is enhanced by the trust region mechanism [28] and the functions involved are sufficiently smooth.

Response correction approaches can be roughly categorized into parametric techniques, where the response correction function (2) can be represented using explicit formulas, and non-parametric ones, where the response correction is defined implicitly through certain relations between the coarse and fine models which may not have any compact analytical form. Output space mapping [11], [12], manifold mapping [22], and multi-point response correction [23] belong to the first group. We describe them briefly in Section 3. The second group includes adaptive response correction [24] and shape-preserving response prediction [25] and is presented in detail in Section 4.

3. Parametric Response Correction Techniques for Microwave Optimization

In this section, we briefly review several parametric response correction techniques that have been recently developed and applied in simulation-driven microwave optimization. More specifically, we discuss several variations of output space mapping (OSM) [12], manifold mapping (MM) [22], as well as mention multi-point response correction [23].

The simplest realization of the response correction is output SM [11], where

$$\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x}) + \boldsymbol{d}^{(i)} \tag{4}$$

Here, $\boldsymbol{d}^{(i)} = \boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i)})$, which ensures zero-order consistency. Note that this way of creating the surrogate model is only based on the fine model response at the most current iteration point $\boldsymbol{x}^{(i)}$. Typically, output SM is used as a supplement for other SM techniques [15]. Provided that the coarse model is sufficiently accurate, output SM can, also, work as a stand-alone technique [15].

The model (4) can be enhanced as follows:

$$\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x}) + \boldsymbol{d}^{(i)} + \boldsymbol{E}^{(i)} \cdot (\boldsymbol{x} - \boldsymbol{x}^{(i)}) \tag{5}$$

where $\boldsymbol{d}^{(i)}$ is as in (4) and $\boldsymbol{E}^{(i)} = \boldsymbol{J}[\boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i)})]$, which ensures first-order consistency. Alternatively, Jacobian of $\boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i)})$ can be replaced by a suitable approximation obtained using, e.g., Broyden update [29]. As before, the model (5) is normally used as a supplement but it can, also, work as a stand-alone technique, particularly when the algorithm (3) is enhanced by a trust-region approach [28], [30].

The response correction model can be defined by explicitly using most of the available fine model data. According to the manifold mapping (MM) approach [22], the surrogate model is defined as

$$\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{R}_f(\boldsymbol{x}^{(i)}) + \boldsymbol{S}^{(i)}\left(\boldsymbol{R}_c(\boldsymbol{x}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i)})\right) \tag{6}$$

with $\boldsymbol{S}^{(i)}$ being the $m \times m$ correction matrix $\boldsymbol{S}^{(i)} = \boldsymbol{\Delta F} \cdot \boldsymbol{\Delta C}^{\dagger}$ where $\boldsymbol{\Delta F} = [\boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_f(\boldsymbol{x}^{(i-1)}) \ \ldots \ \boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_f(\boldsymbol{x}^{(\max\{i-n,0\})})]$ and $\boldsymbol{\Delta C} = [\boldsymbol{R}_c(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i-1)}) \ \ldots \ \boldsymbol{R}_c(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(\max\{i-n,0\})})]$. The pseudoinverse, denoted by †, is defined as $\boldsymbol{\Delta C} = \boldsymbol{V}_{\boldsymbol{\Delta C}}(\boldsymbol{\Sigma}_{\boldsymbol{\Delta C}})^{\dagger}(\boldsymbol{U}_{\boldsymbol{\Delta C}})^{T}$, where $\boldsymbol{U}_{\boldsymbol{\Delta C}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\Delta C}}$, and $\boldsymbol{V}_{\boldsymbol{\Delta C}}$ are the factors in the singular value decomposition of $\boldsymbol{\Delta C}$. The matrix $\boldsymbol{\Sigma}_{\boldsymbol{\Delta C}}^{\dagger}$ is the result of inverting the nonzero entries in $\boldsymbol{\Sigma}_{\boldsymbol{\Delta C}}$, leaving the zeroes invariant [22]. Although MM does not explicitly use sensitivity information, the surrogate and the fine model Jacobians become more and more similar to each other towards the end of the MM optimization process (i.e., when $\|\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(i-1)}\| \to 0$) so that the surrogate (approximately) satisfies both zero- and first-order consistency conditions [21] with $\boldsymbol{R}_f$. This allows for a more accurate location of the fine model optimum.

Another way of generalizing output SM in order to exploit as much of available fine model data as possible is multi-point response correction [23]. We omit the formulation details here for the sake of brevity.

4. Exploiting Maximum Knowledge: Non-Parametric Response Correction Techniques

In this section, we describe non-parametric response correction techniques. Here, the surrogate model is defined implicitly through certain relations between the coarse and fine models which may not have any compact analytical form. On the other hand, non-parametric correction is able—in many cases—to better exploit the knowledge about the system embedded in the coarse model than the parametric one. We discuss the following approaches: adaptive response correction [24], and shape-preserving response prediction [25].

4.1. Adaptive Response Correction

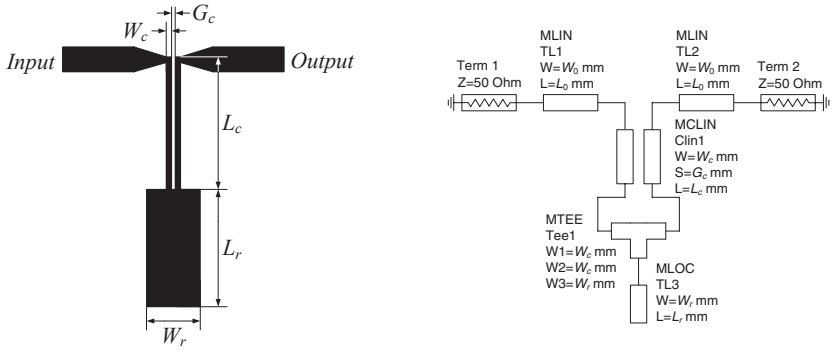
Adaptive response correction (ARC) is a generalization of output SM, which makes the correction term design-variable-dependent so that we have

$$\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x}) + \boldsymbol{\Delta}_r(\boldsymbol{x}, \boldsymbol{x}^{(i)}) \tag{7}$$

where $\boldsymbol{\Delta}_r(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(i)}) = \boldsymbol{d}^{(i)} = \boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i)})$ i.e., a zero-order consistency condition is satisfied at $\boldsymbol{x}^{(i)}$. A characteristic feature of ARC is that the correction term is determined in such a way that this modification reflects the changes of $\boldsymbol{R}_c$ during the process of surrogate model optimization. In particular, if the response of $\boldsymbol{R}_c$ shifts or changes its shape with respect to a free parameter (typically, the signal frequency), the response correction term should track these changes.

The concept of ARC is best explained by an example. Consider a wideband bandstop microstrip filter [31] shown in Fig. 1. Figure 2 shows the fine and coarse (in fact, we use a space-mapped coarse model response here [24]) at two different designs, as well as the corresponding output SM terms $\boldsymbol{d}$. It can be observed that the (frequency-wise) changes of the output SM correction term correspond to the changes of the coarse/fine model responses. These changes can be tracked by performing suitable frequency scaling of the coarse model response and then applied to the original correction term. This is illustrated in Fig. 3, where the original output SM correction term (at certain design $\boldsymbol{x}^{(0)}$), as well as the predicted one (at other design $\boldsymbol{x}^{(1)}$) obtained using the frequency scaling function determined by tracking the coarse model response changes, are shown. A rigorous formulation of ARC can be found in [24]; it is omitted here for the sake of brevity. Figure 4 shows the results of optimizing the surrogate model using the output SM correction (Fig. 4(a)) and ARC (Fig. 4(b)). The optimized surrogate model design in Fig. 4(a) is not as good as that in Fig. 4(b) because the constant output SM correction terms causes a significant distortion of the model response while moving away from $\boldsymbol{x}^{(0)}$.

Table 1 shows the results of optimizing the bandstop filter in Fig. 1 using the surrogate model based on output SM and ARC. The latter technique is able to yield a better result using only three fine model evaluations, whereas the optimization cost with output SM is eight fine model evaluations. Figure 5 shows the fine model response at the final design obtained with ARC.



(a) (b)

Fig. 1. Wideband bandstop filter [31]: (a) geometry, (b) coarse model (Agilent ADS).

Table 1: Optimization results for the wideband bandstop microstrip filter

Surrogate Model	Final Specification Error [dB]	Number of Fine Model Evaluations*
Standard Output Space Mapping	–1.1	8
Adaptive Response Correction	–2.1	3

* Excludes the fine model evaluation at the starting point.

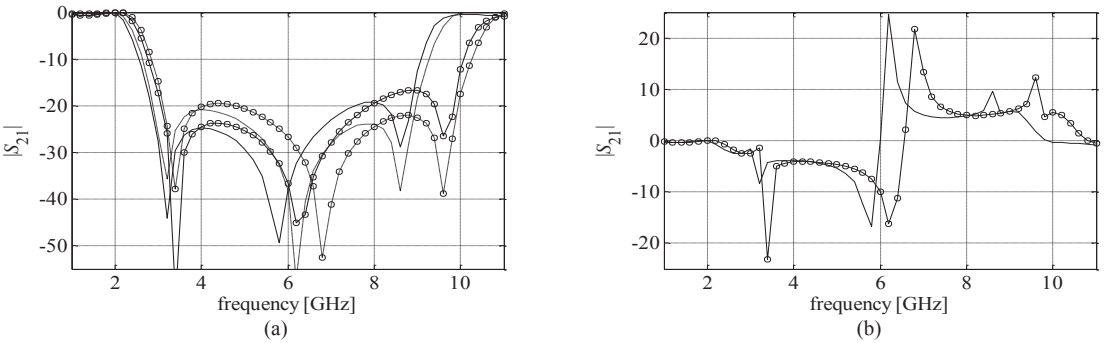


Fig. 2. Wideband bandstop filter: (a) fine (solid line) and coarse model (dashed line) responses at certain design, as well as the fine (solid line with circle markers) and coarse model (dashed line with circle markers) responses at a different design; (b) standard output SM correction terms corresponding to responses shown in Fig. 2.

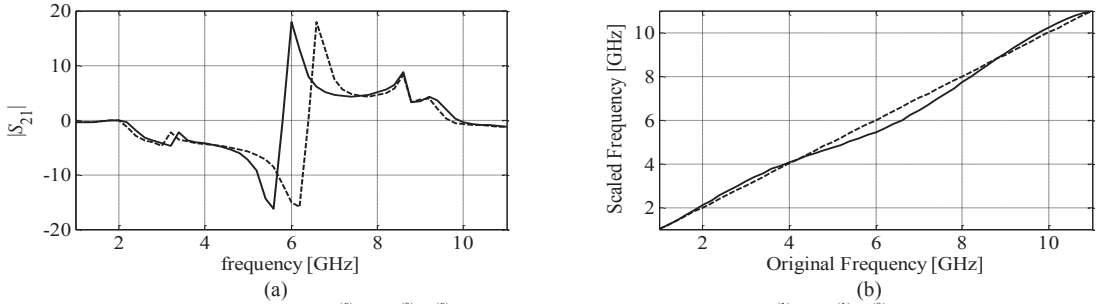


Fig. 3. (a) Correction terms at certain design $\boldsymbol{x}^{(0)}$, $\Delta_r(\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(0)})$, (solid line), and at the other design $\boldsymbol{x}^{(1)}$, $\Delta_r(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(0)})$ (dashed line). Plots obtained for the adaptive response correction method; (b) The frequency scaling function used to obtain the dashed line in Fig. 4 (solid line). The scaling accounts for the changes (in frequency) of the coarse model response while going from $\boldsymbol{x}^{(0)}$ to $\boldsymbol{x}^{(1)}$. The identity function plot is shown as dashed line.

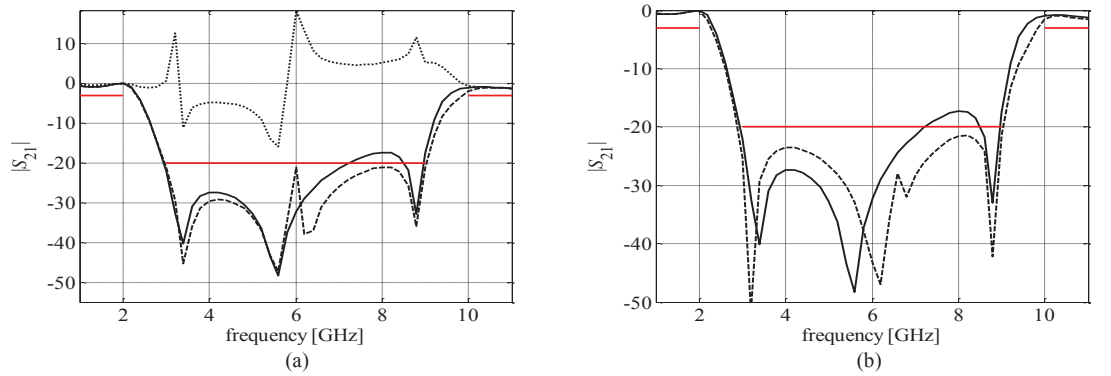


Fig. 4. Wideband bandstop filter: (a) initial (solid line) and optimized (dashed line) surrogate model response when output SM is used to create the surrogate; (b) initial (solid line) and optimized (dashed line) surrogate model response when ARC is used to create the surrogate. The plots in (b) correspond to $\boldsymbol{x}^{(0)}$ and $\boldsymbol{x}^{(1)}$ in Fig. 5.

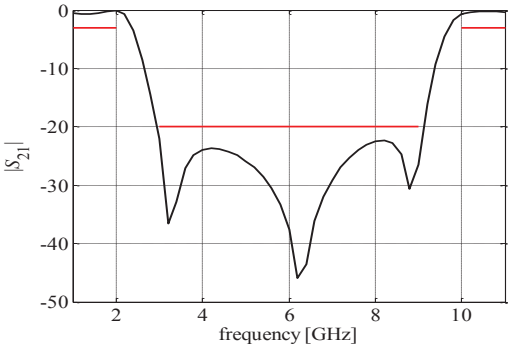


Fig. 5. Wideband bandstop filter: fine model response at the design found by the adaptive response correction technique.

4.2. Shape-Preserving Response Prediction

Shape-preserving response prediction (SPRP) [25] is one of the most recent non-parametric response correction techniques. SPRP constructs the surrogate model assuming that the change of the fine model response due to the adjustment of the design variables can be predicted using the actual changes of the coarse model response. Therefore, it is critically important that the coarse model is physically based, which ensures that the effect of the design parameter variations on the model response is similar for both the fine and coarse models. The change of the coarse model response is described by the translation vectors corresponding to certain (finite) number of characteristic points of the model's response. These translation vectors are subsequently used to predict the change of the fine model response with the actual response of $\boldsymbol{R}_f$ at the current iteration point, $\boldsymbol{R}_f(\boldsymbol{x}^{(i)})$, used as a reference.

Figure 6(a) shows an example coarse model response, $|S_{21}|$ in the frequency range 8 GHz to 18 GHz, at the design $\boldsymbol{x}^{(i)}$, as well as a coarse model response at some other design $\boldsymbol{x}$. The responses come from the double folded stub bandstop filter example considered in [25]. Circles denote characteristic points of $\boldsymbol{R}_c(\boldsymbol{x}^{(i)})$, here, selected to represent $|S_{21}| = –3$ dB, $|S_{21}| = –20$ dB, and the local $|S_{21}|$ maximum (at about 13 GHz). Squares denote corresponding characteristic points for $\boldsymbol{R}_c(\boldsymbol{x})$, while line segments represent the translation vectors ("shift") of the characteristic points of $\boldsymbol{R}_c$ when changing the design variables from $\boldsymbol{x}^{(i)}$ to $\boldsymbol{x}$. Because the coarse model is physically-based, the fine model response at a given design, here, $\boldsymbol{x}$, can be predicted using the same translation vectors applied to the corresponding characteristic points of the fine model response at $\boldsymbol{x}^{(i)}$, $\boldsymbol{R}_f(\boldsymbol{x}^{(i)})$. This is illustrated in Fig. 6(b). Note that SPRP surrogate model does not explicitly use any parameters, which makes it easy to implement. A rigorous formulation of SPRP can be found in [25].

As an illustration, consider the dual-band bandpass filter [32] (Fig. 7(a)). The design parameters are $\boldsymbol{x} = [L_1\ L_2\ S_1\ S_2\ S_3\ d\ g\ W]^T$ mm. The fine model is simulated in Sonnet ***em*** [33]. The design specifications are $|S_{21}| \geq –3$ dB for $0.85\ \text{GHz} \leq \omega \leq 0.95\ \text{GHz}$ and $1.75\ \text{GHz} \leq \omega \leq 1.85\ \text{GHz}$, and $|S_{21}| \leq –20$ dB for $0.5\ \text{GHz} \leq \omega \leq 0.7\ \text{GHz}$, $1.1\ \text{GHz} \leq \omega \leq 1.6\ \text{GHz}$ and $2.0\ \text{GHz} \leq \omega \leq 2.2\ \text{GHz}$. The coarse model is implemented in Agilent ADS [34] (Fig. 7(b)). The initial design is $\boldsymbol{x}^{(0)} = [16.14\ 17.28\ 1.16\ 0.38\ 1.18\ 0.98\ 0.98\ 0.20]^T$ mm (the optimal solution of $\boldsymbol{R}_c$). The following characteristic points are selected to set up the SPRP model: four points for which $|S_{21}| = –20$ dB, four points with $|S_{21}| = –5$ dB, as well as 6 additional points located between –5 dB points.

For the purpose of optimization, the coarse model was enhanced by tuning the dielectric constants and the substrate heights of the microstrip models corresponding to the design variables L_1, L_2, d and g (original values of ε_r and H were 10.2 and 0.635 mm, respectively). The filter was optimized using shape-preserving response prediction, space mapping and adaptive response correction.

Table 2 shows the optimization results. SPRP was compared to output SM as well as ARC. All of the considered methods were able to yield solutions satisfying the design specifications, which is mostly because of using the enhanced coarse model. However, the quality of the designs produced by SPRP better than the quality the designs yielded by other methods. Also, computational cost of SPRP is lower than that of other approaches, particularly output SM. Figure 8 shows the initial fine model response as well as the fine model response at the design obtained using SPRP.

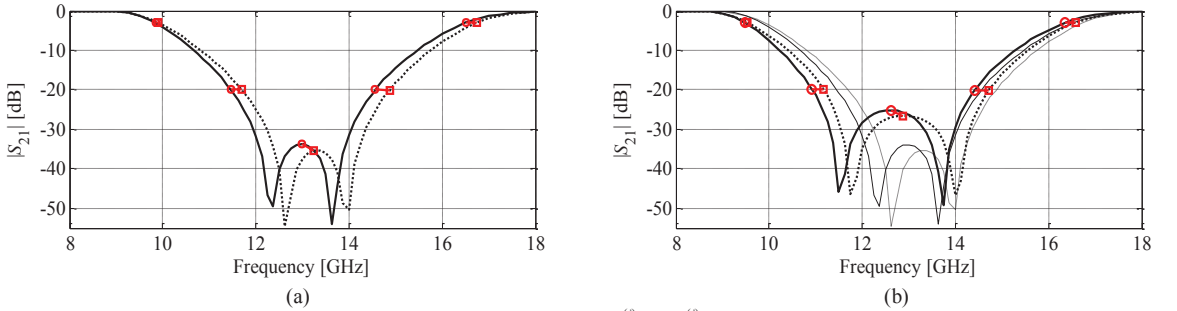
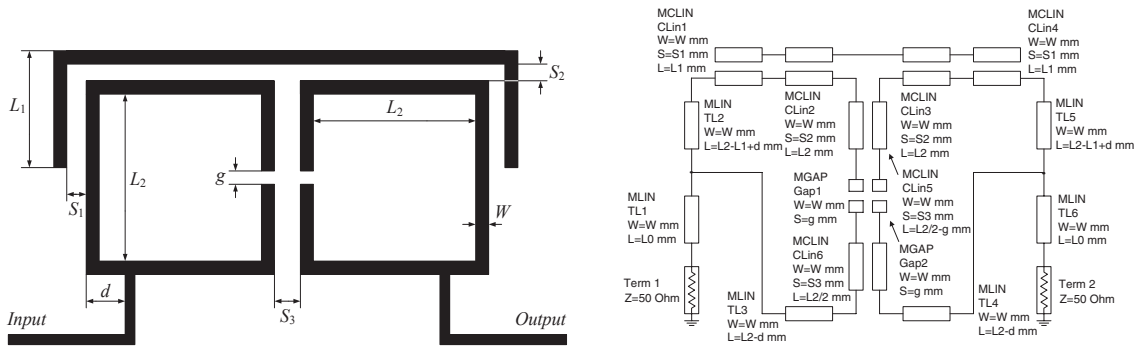


Fig. 6. SPRP concept: (a) Example coarse model response at the design $\boldsymbol{x}^{(i)}$, $\boldsymbol{R}_c(\boldsymbol{x}^{(i)})$ (solid line), the coarse model response at $\boldsymbol{x}$, $\boldsymbol{R}_c(\boldsymbol{x})$ (dotted line), characteristic points of $\boldsymbol{R}_c(\boldsymbol{x}^{(i)})$ (circles) and $\boldsymbol{R}_c(\boldsymbol{x})$ (squares), and the translation vectors (short lines); (b) Fine model response at $\boldsymbol{x}^{(i)}$, $\boldsymbol{R}_f(\boldsymbol{x}^{(i)})$ (solid line) and the predicted fine model response at $\boldsymbol{x}$ (dotted line) obtained using SPRP based on characteristic points of Fig. 6(a); characteristic points of $\boldsymbol{R}_f(\boldsymbol{x}^{(i)})$ (circles) and the translation vectors (short lines) were used to find the characteristic points (squares) of the predicted fine model response; coarse model responses $\boldsymbol{R}_c(\boldsymbol{x}^{(i)})$ and $\boldsymbol{R}_c(\boldsymbol{x})$ are plotted using thin solid and dotted line, respectively [25].



(a) (b)

Fig. 7. Dual-band bandpass filter: (a) geometry [32], (b) coarse model (Agilent ADS).

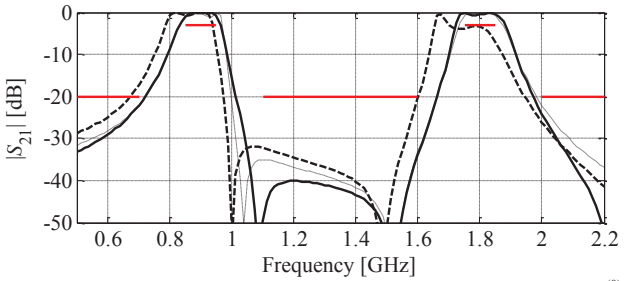


Fig. 8. Dual-band bandpass filter: fine model (dashed line) and coarse model (thin dashed line) response at $\boldsymbol{x}^{(0)}$, and the optimized fine model response (solid line) at the design obtained using shape-preserving response prediction.

Table 2: Optimization results for dual-band bandpass filter

Algorithm	Final Specification Error [dB]	Number of Fine Model Evaluations[1]
Output SM[2]	–1.4	8
Output SM[2] + ISM[3]	–1.9	8
ARC [24]	–1.8	3
ARC [24] + ISM[3]	–1.5	4
SPRP	–2.0	3
SPRP + ISM[3]	–1.9	2

[1] Excludes the fine model evaluation at the starting point.
[2] The surrogate model is of the form $\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x}) + [\boldsymbol{R}_f(\boldsymbol{x}^{(i)}) - \boldsymbol{R}_c(\boldsymbol{x}^{(i)})]$.
[3] The surrogate model is of the form $\boldsymbol{R}_s^{(i)}(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x} + \boldsymbol{c}^{(i)})$; $\boldsymbol{c}^{(i)}$ found using parameter extraction [11].

5. Exploiting Maximum Knowledge: Adaptively Adjusted Design Specifications

It is not necessary to remove the discrepancies between the low- and high-fidelity models by correcting the low-fidelity model. Another way is to "absorb" the model misalignment by proper adjustment of the design specifications. In microwave engineering, most of the design tasks can be formulated as minimax problems with upper and lower specifications and it is easy to implement modifications by, for example, shifting the specification levels, corresponding frequency bands. This approach, both easy to implement and efficient, is exploited by adaptively adjusted design specifications (AADS) technique [26] described in this section. AADS consists of the following two simple steps that can be iterated if necessary:

1. Modify the original design specifications in order to take into account the difference between the responses of $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ at their characteristic points.
2. Obtain a new design by optimizing the coarse model with respect to the modified specifications.

Characteristic points of the responses should correspond to the design specification levels. They should also include local maxima/minima of the respective responses at which the specifications may not be satisfied. Figure 9(b) shows characteristic points of $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ for our bandstop filter example. The points correspond to –3 dB and –30 dB levels as well to the local maxima of the responses. As one can observe in Fig. 9(b) the selection of points is rather straightforward.

In the first step of AADS optimization procedure, the design specifications are modified (or mapped) so that the level of satisfying/violating the modified specifications by the coarse model response corresponds to the satisfaction/violation

levels of the original specifications by the fine model response. More specifically, for each edge of the specification line, the edge frequency is shifted by the difference of the frequencies of the corresponding characteristic points, e.g., the left edge of the specification line of –30 dB is moved to the right by about 0.7 GHz, which is equal to the length of the line connecting the corresponding characteristic points in Fig. 9(b). Similarly, the specification levels are shifted by the difference between the local maxima/minima values for the respective points, e.g., the –30 dB level is shifted down by about 8.5 dB because of the difference of the local maxima of the corresponding characteristic points of $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$. Modified design specifications are shown in Fig. 9(c). The coarse model is subsequently optimized with respect to the modified specifications and the new design obtained this way is treated as an approximated solution to the original design problem (i.e., optimization of the fine model with respect to the original specifications). Steps 1 and 2 can be repeated if necessary. As demonstrated later, substantial design improvement is typically observed after the first iteration, however, additional iterations may bring further enhancement.

An important prerequisite of AADS is that the coarse model is physics-based, in particular, the adjustment of the design variables has similar effect on the response for both $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$. In such a case the coarse model design that is obtained in the second stage of the AADS procedure (i.e., optimal with respect to the modified specifications) will be (almost) optimal for $\boldsymbol{R}_f$ with respect to the original specifications. As shown in Fig. 9, the absolute matching between the models is not as important as the shape similarity. If the similarity between the fine and coarse model response is not sufficient the AADS technique may not work well. A generalized version of AADS that alleviates this difficulty can be found in [26].

As an illustration that demonstrate operation and efficiency of AADS, consider the second example, consider the bandpass microstrip filter with open stub inverter [35] (Fig. 10(a)). The design parameters are $\boldsymbol{x} = [L_1\ L_2\ L_3\ S_1\ S_2\ W_1]^T$. The fine model is simulated in FEKO [36]. The design specifications are $|S_{21}| \le$ –20dB for 1.5GHz $\le \omega \le$ 1.8GHz, $|S_{21}| \ge$ –3dB for 1.95GHz $\le \omega \le$ 2.05GHz and $|S_{21}| \le$ –20dB for 2.2GHz $\le \omega \le$ 2.5GHz. The coarse model is implemented in Agilent ADS [34] (Fig. 10(b)). The initial design is the coarse model optimal solution $\boldsymbol{x}^{(0)} = [25.00\ 5.00\ 1.221\ 0.652\ 0.187\ 0.100]^T$ mm (specification error +15.7 dB).

The first iteration of AADS already yields a design satisfying the specifications, $\boldsymbol{x}^{(1)} = [23.79\ 5.00\ 1.00\ 0.694\ 0.192\ 0.10]^T$ mm (specification error –0.6 dB). After the second iteration, the design was further improved to $\boldsymbol{x}^{(2)} = [23.68\ 5.00\ 1.00\ 0.717\ 0.193\ 0.10]^T$ mm (specification error –1.7 dB). Figure 11 shows the fine and coarse model responses at $\boldsymbol{x}^{(0)}$ and the fine model response at the final design. For the sake of comparison, the filter was also optimized using the frequency SM algorithm. The design obtained in three iterations, $[23.66\ 5.00\ 1.00\ 0.654\ 0.188\ 0.100]^T$ mm, satisfies the design specifications, however, it is not as good at the one obtained using the procedure proposed in this work (specification error –0.8 dB).

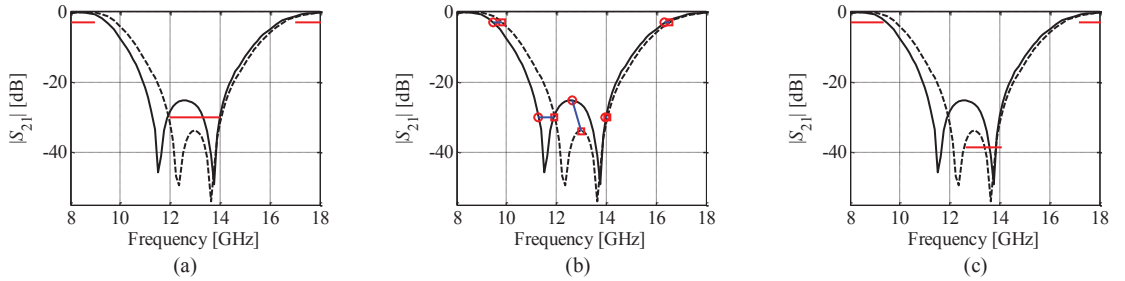


Fig. 9. Bandstop filter example (responses of $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ are marked with solid and dashed line, respectively): (a) responses of $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ at the initial design (optimum of $\boldsymbol{R}_c$) as well as the original design specifications, (b) characteristic points of the responses corresponding to the specification levels (here, –3 dB and –30 dB) and to the local response maxima, (c) responses of $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ at the initial design and the modified design specifications.

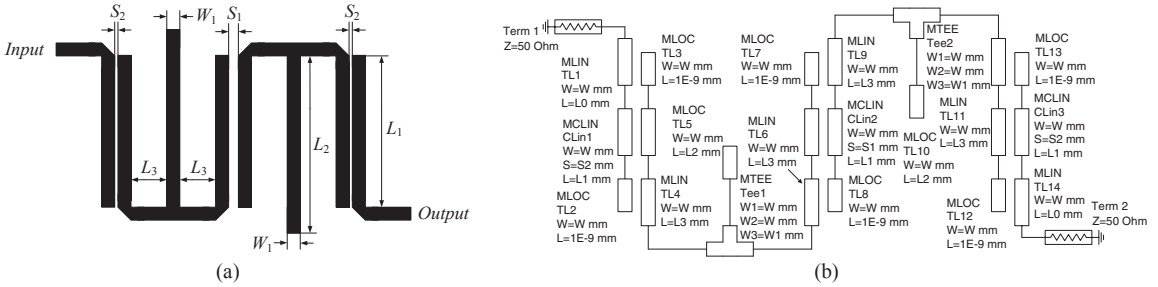


Fig. 10. Bandpass filter with open stub inverter: (a) geometry [35], (b) coarse model (Agilent ADS).

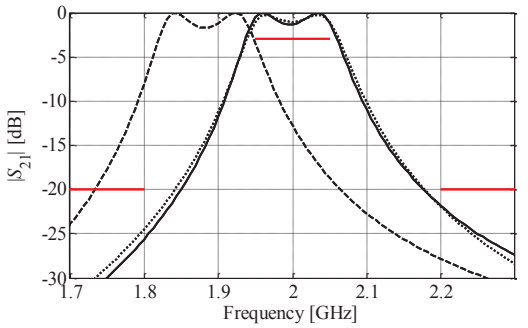


Fig. 11. Bandpass filter with open stub inverter: $\boldsymbol{R}_f$ response (solid line) at the final design obtained after two iterations of our optimization procedure; fine (dashed line) and coarse (dotted line) model responses at the initial design.

6. Discussion and Recommendations

Here, we attempt to give a qualitative comparison of knowledge-based design techniques presented in this paper, as well as formulate some guidelines and recommendations for the readers interested in applying these methods in their research and/or design work. The main factors are the complexity of implementation, computational efficiency, robustness, as well as the range of applications.

In terms of implementation, the parametric response correction techniques are very straightforward. Output space mapping, manifold mapping and multi-point response correction construct the surrogate model by using analytical formulas, therefore, they are easy to implement. Non-parametric methods may be more involved because the surrogate models are constructed by considering some auxiliary quantities such as scaling functions (adaptive response correction) or characteristic points and translation vectors (SPRP), which generally depend on the response shape. Also, in SPRP, the user is responsible for defining the characteristic points on case-to-case basis. On the other hand, the AADS approach described in Section 5 is probably the easiest to implement out of all the methods considered in this work.

In terms of computational complexity, non-parametric methods prove to be more efficient. The reason is that both ARC and SPRP, as well as AADS, exploit the knowledge embedded in the low-fidelity model to a larger extent than the parametric methods which only do a local model alignment. As indicated by the examples, ARC and SPRP are capable of yielding a satisfactory design after two or three iterations, whereas output SM or MM typically require more iterations. On the other hand, parametric techniques tend to be more robust, particularly when embedded in the trust-region framework which improves their convergence properties. Also, the use of (even approximated) fine model sensitivity improves their ability to locate the fine model optimum accurately. Parametric techniques are also more generic than ARC, SPRP and AADS. For the latter, some considerations are necessary in order to select a proper realization of the scaling function (ARC), the definition of the characteristic points (SPRP), or the analysis of the model responses to properly modify design specifications (AADS). Also, the SPRP technique assumes one-to-one correspondence between characteristic points of the coarse and fine models at all considered design, which may be difficult to satisfy for certain problems.

Based on the above remarks, parametric methods, particularly output SM, are recommended for less experienced users, and when the underlying coarse model is relatively fast (e.g., equivalent circuit). More experienced users are encouraged to try either ARC or SPRP, particularly if the available coarse model is relatively expensive (e.g., obtained from coarse-discretization EM simulation). AADS, because of its simplicity, can be readily used by both experienced and novice users; in general, it tends to be more reliable when the coarse model is more accurate (e.g., obtained through coarse-discretization EM simulation). In either case, the quality of the coarse model is one of the key factors, therefore its preconditioning using any space mapping transformation is recommended.

References

1. Wu, K., “Substrate Integrated Circuits (SiCs) – A new paradigm for future Ghz and Thz electronic and photonic systems,” *IEEE Circuits and Systems Society Newsletter*, vol. 3, p. 1, 2009.
2. Schantz, H. *The art and science of ultrawideband antennas*, Artech House, 2005.
3. Petosa, A., *Dielectric Resonator Antenna Handbook*, Artech House, 2007.

4. Nikolova, N.K., Li, Y., Li, Y., and Bakr, M.H., "Sensitivity analysis of scattering parameters with electromagnetic time-domain simulators," *IEEE Trans. Microwave Theory Tech*., vol. 54, pp. 1598-1610, 2006.
5. CST Microwave Studio, CST AG, Bad Nauheimer Str. 19, D-64289 Darmstadt, Germany, 2011.
6. HFSS, release 13.0, ANSYS, 2010.
7. Kabir, H., Wang, Y., Yu, M., and Zhang, Q.J., "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microwave Theory Tech.*, vol. 56, pp. 867-879, 2008.
8. Zhang, L., Zhang, Q. J., and Wood, J., "Statistical neuro-space mapping technique for large-signal modeling of nonlinear devices," *IEEE Trans. Microwave Theory Tech.*, vol. 56, pp. 2453-2467, 2008.
9. Miraftab, V., and Mansour, R.R., "EM-based microwave circuit design using fuzzy logic techniques," *IEE Proc. Microwaves, Antennas & Propagation*, vol. 153, pp. 495-501, 2006.
10. Couckuyt, I., Declercq, F., Dhaene, T., Rogier, H., and Knockaert, L., "Surrogate-based infill optimization applied to electromagnetic problems," *Int. J. RF and Microwave CAE*, vol. 20, pp. 492-501, 2010.
11. Bandler, J.W., Cheng, Q.S., Dakroury, S.A., Mohamed, A.S., Bakr, M.H., Madsen, K., and Sondergaard, J., "Space mapping: the state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 337-361, 2004.
12. Koziel S., Bandler J.W., and Madsen, K., "A space mapping framework for engineering optimization: theory and implementation," *IEEE Trans. Microwave Theory Tech*., vol. 54, pp. 3721-3730, 2006.
13. Amari, S., LeDrew, C., and Menzel, W., "Space-mapping optimization of planar coupled-resonator microwave filters," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 2153-2159, 2006.
14. Ismail, M.A., Smith, D., Panariello, A., Wang, Y., and Yu, M., "EM-based design of large-scale dielectric-resonator filters and multiplexers by space mapping," *IEEE Trans. Microwave Theory Tech.,* vol. 52, pp. 386-392, 2004.
15. Koziel, S., Cheng, Q.S., and Bandler, J.W., "Space mapping," *IEEE Microwave Magazine*, vol. 9, pp. 105-122, 2008.
16. Quyang, J., Yang, F., Zhou, H., Nie, Z., and Zhao, Z., "Conformal antenna optimization with space mapping," *J. of Electromagn. Waves and Appl.*, vol. 24, pp. 251-260, 2010.
17. Rautio, J.C., "Perfectly calibrated internal ports in EM analysis of planar circuits," *IEEE MTT-S Int. Microwave Symp. Dig.*, Atlanta, GA, pp. 1373-1376, 2008.
18. Swanson, D., and Macchiarella, G., "Microwave filter design by synthesis and optimization," *IEEE Microwave Magazine*, 8, pp. 55-69, 2007.
19. Cheng, Q.S., Rautio, J.C., Bandler, J.W., and Koziel, S., "Progress in simulator-based tuning—the art of tuning space mapping," *IEEE Microwave Magazine*, vol. 11, pp. 96-110, 2010.
20. Koziel, S., Meng, J., Bandler, J.W., Bakr, M.H., and Cheng, Q.S., "Accelerated microwave design optimization with tuning space mapping," *IEEE Trans. Microwave Theory and Tech.*, vol. 57, pp. 383-394, 2009.
21. Alexandrov, N.M., and Lewis, R.M., "An overview of first-order model management for engineering optimization," *Optimization and Engineering*, vol. 2, pp. 413-430, 2001.
22. Echeverria, D., and Hemker, P.W., "Space mapping and defect correction," *CMAM The International Mathematical Journal Computational Methods in Applied Mathematics*, vol. 5, pp.107-136, 2005.
23. Koziel, S., "Improved microwave circuit design using multipoint-response-correction space mapping and trust regions," *Int. Symp. Antenna Technology and Applied Electromagnetics*, *ANTEM 2010*, Ottawa, Canada, 2010.
24. Koziel, S., Bandler, J.W., and Madsen, K., "Space mapping with adaptive response correction for microwave design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 57, pp. 478-486, 2009.
25. Koziel, S., "Shape-preserving response prediction for microwave design optimization," *IEEE Trans. Microwave Theory and Tech.*, vol. 58, pp. 2829-2837, 2010.
26. Koziel, S., "Adaptively adjusted design specifications for efficient optimization of microwave structures," *Progress in Electromagnetic Research B (PIER B)*, vol. 21, pp. 219-234, 2010.
27. Alexandrov, N.M., Dennis, J.E., Lewis, R.M., and Torczon, V., "A trust region framework for managing use of approximation models in optimization," *Struct. Multidisciplinary Optim.*, vol. 15, pp. 16-23, 1998.
28. Conn, A.R., Gould, N.I.M., and Toint, P.L., *Trust Region Methods*, MPS-SIAM Series on Optimization, 2000.
29. Broyden, C.G., "A class of methods for solving nonlinear simultaneous equations," *Math. Comp.*, vol. 19, pp. 577-593, 1965.
30. Koziel, S., Bandler, J.W., and Cheng, Q.S., "Robust trust-region space-mapping algorithms for microwave design optimization," *IEEE Trans. Microwave Theory and Tech.*, vol. 58, pp. 2166-2174, 2010.
31. Hsieh, M.Y., and Wang, S.M., "Compact and wideband microstrip bandstop filter," *IEEE Microwave and Wireless Component Letters*, vol. 15, pp. 472-474, 2005.
32. Guan, X., Ma, Z., Cai, P., Anada, T., and Hagiwara, G., "A microstrip dual-band bandpass filter with reduced size and improved stopband characteristics," *Microwave and Opt. Tech. Lett.*, vol. 50, pp. 618-620, 2008.
33. Sonnet Software, Inc., ***em***™ Version 12.54, 100 Elwood Davis Road, North Syracuse, NY 13212, USA, 2010.
34. Agilent ADS, Agilent Technologies, 1400 Fountaingrove Parkway, Santa Rosa, CA 95403-1799, 2008.
35. Lee, J.R., Cho, J.H., and Yun, S.W., "New compact bandpass filter using microstrip $\lambda/4$ resonators with open stub inverter," *IEEE Microwave and Guided Wave Letters*, vol. 10, pp. 526-527, 2000.
36. FEKO, Suite 6.0, EM Software & Systems-S.A. (Pty) Ltd, 32 Techno Lane, Technopark, Stellenbosch, 7600, South Africa, 2010.
37. Bandler, J.W., Cheng, Q.S., Nikolova, N.K., and Ismail, M.A., "Implicit space mapping optimization exploiting preassigned parameters," *IEEE Trans. Microwave Theory Tech*., vol. 52, pp. 378-385, 2004.
38. Rayas-Sánchez, J. E., and Gutiérrez-Ayala, V., "EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 4528-4537, 2006.